

引用格式:肖擎曜,胡非,范绍佳.风能数值预报的模式输出统计(MOS)研究[J].资源科学,2017,39(1):116-124. [Xiao Q Y, Hu F, Fan S J. Model output statistics and wind power numerical prediction[J]. *Resources Science*, 2017, 39(1): 116-124]. DOI: 10.18402/resci.2017.01.12

风能数值预报的模式输出统计(MOS)研究

肖擎曜^{1,2}, 胡非¹, 范绍佳³

(1. 中国科学院大气物理研究所, LAPC, 北京 100029; 2. 中国科学院大学, 北京 100049;

3. 中山大学大气科学院, 广州 510275)

摘要:利用中尺度气象模式WRF输出的风速、温度、摩擦速度等变量作为待选因子,在其中选择合适的因子与风电场的实测风速建立多元线性模式输出统计(MOS)模型,对陕西省延长县风电场的风速进行预报。详细分析了因子数量、模型训练长度对预报结果的影响。研究表明:①所选因子不需要太多,推荐只选两个;②保证有600或以上时次的历史资料进行训练;③满足以上条件,MOS预报的均方根误差比单纯的模式预报误差平均可减少0.23 m/s。该方法实现简单,计算快捷,可实际应用于风电预报中。

关键词:模式输出统计;多元线性回归;风电功率预报;WRF

DOI: 10.18402/resci.2017.01.12

1 引言

当前,中国面临着能源短缺以及环境问题所带来的挑战,如何开发利用可再生能源,是关系到可持续发展的战略问题。风能具有无污染、生态影响小、适合大规模开发等优点,是一种优秀的可再生能源。但与其他发电方式,如火电、水电或核电不同,风力发电由于风的波动性和间歇性,不可避免地导致输出功率的不稳定,对电能质量、电网稳定性和保护装置影响极大。因此,及时准确的风电功率预报对提高电网对风电的接纳能力及风能资源的合理利用开发起到关键作用^[1]。

风速预报是风电功率预报的基础。风电场风速预报方法一般有两类,统计方法和物理方法。统计方法主要包括自回归滑动平均(ARMA)^[2,3]、卡尔曼滤波(Kalman filter)^[4]、支撑向量机(SVM)^[5,6]、人工神经网络(ANN)^[7,8]等。基本思路为利用风电场历史风速数据,建立统计模型,直接预报。物理方法利用全球模式和中尺度模式,计算出风电场附近的风速、气温等资料^[9]。但由于风电场风速预报对时

空尺度的要求十分精细,且风电场一般处于山梁等复杂地形,直接利用中尺度模式进行计算,往往无法满足风电场风速预报的精度要求。对此,需要再用动力CFD方法^[10,11]或统计方法在模式输出量的基础上进行处理,才能得到符合风力发电要求的风速预报^[12]。对于风速预报的模式统计后处理方法,主要有模式输出统计、卡尔曼滤波^[13]、BP神经网络^[14]以及自适应偏最小二乘法^[15]等。其中,使用最广泛的是模式输出统计(MOS)方法。

Glahn等首次提出了MOS应用于气象要素预报,即将数值预报的历史值与预报要素的历史实况建立统计关系,预报的时候代入模式的预报因子,对模式的系统误差有明显的订正能力^[16]。Landberg指出,利用流体力学方法对模式进行后处理,其结果很大程度也是线性的,为利用线性回归进行MOS处理提供了实验支持^[17]。目前,对MOS系统在常规天气预报中的应用已有深入研究并已应用于实际。国家气象中心刘还珠等建立了MOS预报系统,为日常的短期要素预报提供参考^[18]。陈豫英等利用

收稿日期:2016-09-10;修订日期:2016-12-05

基金项目:国家自然科学基金项目(11472272)。

作者简介:肖擎曜,男,广东广州人,博士生,主要从事大气湍流及风能应用研究。E-mail: xiaoqy@mail.iap.ac.cn

通讯作者:胡非, E-mail: hufei@mail.iap.ac.cn

2017年1月

MOS进行风的精细化预报,得到结果比原MM5模式预报水平有显著提高^[19]。林行等利用MOS技术为2008年奥帆赛提供了风场预报^[20]。Ranaboldo将MOS技术应用于风电预报领域,用模式输出的各层风速、气温及多个湍流参数与风电场风速建立多元线性回归,改善模式预报结果^[21]。

目前在中国MOS技术应用于风电预报的工作还很少。本文在上述工作的基础上,选取陕西省延长县的一个风电场为案例,针对风电场历史数据少、预报时间分辨率要求高(10分钟)的特点,深入探讨了多元线性回归模型的因子选取、训练长度对MOS方法应用于风电场风速预报准确度的影响,并给出了推荐参数。经过多元线性回归的处理后,预报结果与单纯模式预报相比有较大改善。本方法思路清晰,实现方便,计算快捷,有利于广泛应用推广,为提高风电功率预报的准确度提供了一个可选方案,也为风能资源的合理开发利用提供参考。

2 数学方法

模式输出统计(MOS),指在历史模式输出量中选择合适的因子与观测值建立统计关系,在实际预报时将相应的模式输出量代入该统计关系中,得到修正后的预报值。该方法可以有效地减少模式的系统误差。常用多元线性回归模型建立统计关系^[22]。

2.1 多元线性回归模型

现有样本容量为 n 的预报量 $y(t_i)$ 与 p 个因子 $x_1(t_i), x_2(t_i), \dots, x_p(t_i)$,其中 $i=1, 2, \dots, n$ 。假设它们存在如下线性统计关系:

$$y(t_i) = \beta_0 + \beta_1 x_1(t_i) + \beta_2 x_2(t_i) + \dots + \beta_p x_p(t_i) + e(t_i) \quad (1)$$

式中 β 为各个因子的系数; $e(t_i)$ 为相互独立且符合高斯分布的随机变量。

利用这组样本对上述回归模型进行估计得到的估计方程称为多元线性回归方程:

$$\hat{y}(t_i) = b_0 + b_1 x_1(t_i) + b_2 x_2(t_i) + \dots + b_p x_p(t_i) \quad (2)$$

其中 $\hat{y}(t_i)$ 为 $y(t_i)$ 的预报值,而 $b_0, b_1, b_2, \dots, b_p$ 分别为 $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ 的估计,应满足预报量观测值与回归估计值的差值平方和达到最小,即:

$$\begin{aligned} Q &= \sum_{i=1}^n (y(t_i) - \hat{y}(t_i))^2 \\ &= \sum_{i=1}^n (b_0 + b_1 x_1(t_i) + b_2 x_2(t_i) + \dots + b_p x_p(t_i) - y(t_i))^2 \end{aligned}$$

达到最小。

将上式展开,利用最小二乘法可得到系数的标准方程组:

$$\begin{cases} nb_0 + b_1 \sum_{i=1}^n y(t_i) + \dots + b_p \sum_{i=1}^n x_p(t_i) = \sum_{i=1}^n y_i \\ b_0 \sum_{i=1}^n x_1(t_i) + b_1 \sum_{i=1}^n x_1(t_i)^2 + \dots \\ \quad + b_p \sum_{i=1}^n x_1(t_i) x_p(t_i) = \sum_{i=1}^n x_1(t_i) y_i \\ b_0 \sum_{i=1}^n x_2(t_i) + b_1 \sum_{i=1}^n x_2(t_i) x_1(t_i) + \dots \\ \quad + b_p \sum_{i=1}^n x_2(t_i) x_p(t_i) = \sum_{i=1}^n x_2(t_i) y_i \\ b_0 \sum_{i=1}^n x_p(t_i) + b_1 \sum_{i=1}^n x_p(t_i) x_1(t_i) + \dots \\ \quad + b_p \sum_{i=1}^n x_p(t_i)^2 = \sum_{i=1}^n x_p(t_i) y_i \end{cases} \quad (3)$$

解方程组得到 $b_0, b_1, b_2, \dots, b_n$ 。

2.2 因子挑选

回归分析时,需要从可能影响因变量的众多因子中挑选出一批数量适中、关系较好的因子建立回归方程。常用逐步引入法挑选因子。

逐步引入法是在模式给出的 p 个待选因子中,挑选方差贡献最大的,若该因子通过显著性检验,加入回归方程。

第 k 个因子的方差贡献记为:

$$V_k = \frac{b_k^2}{c_{kk}} \quad (4)$$

式中 c_{kk} 为矩阵 $C=(X'X)^{-1}$ 中对角线的第 k 个元素。

X 为带有常数1的因子矩阵:

$$X = \begin{bmatrix} 1 & x_1(t_1) & x_2(t_1) & \dots & x_p(t_1) \\ 1 & x_1(t_2) & x_2(t_2) & \dots & x_p(t_2) \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_1(t_n) & x_2(t_n) & \dots & x_p(t_n) \end{bmatrix} \quad (5)$$

利用统计量 F 检验因子 x_k 是否对风速 v 有显著影响。 F 的表达式为:

$$F = \frac{b_k^2 / c_{kk}}{Q / (n - p - 1)} \quad (6)$$

式中 Q 为残差平方和; n 为样本容量; p 为因子个数。

确定现有因子的回归方程后,把观测值与回归估计值的差值作为预报量,对剩下的待选因子进行

重复因子挑选的步骤,引入新的因子,直到得到满意的回归方程。逐步引入法可以在选出较佳因子方案的前提下,有效地减少运算量。

3 观测资料简介

采用陕西省延长县风电场的70m观测塔测风数据。测风塔位于 $36^{\circ}38'N$, $110^{\circ}15'E$,海拔高度1272m,使用美国NRG测风仪。观测时间为2011年10月18日至2011年11月26日(世界时,下同),观测高度为70m、50m、30m、10m四个高度。时间分辨率为10分钟,共有5760个时次。

利用中尺度气象模式WRF预报该时间段延长县风电场各气象要素。WRF采用3.2版本,区域设置为3层嵌套,双向反馈,以测风塔为中心,水平分辨率分别为27km、9km、3km,网格数分别为 69×69 , 70×70 和 73×64 。垂直分辨率37层,其中1km以下分12层。采用全球分辨率 $1^{\circ} \times 1^{\circ}$,3h一次的GFS/NCEP预报场资料作为初始场和侧边界条件,采用USGS 30s水平分辨率的地形资料。起报时间为每天00:00(世界时),预报时效为24~48h,前24h不做分析是为了给下载全球预报场数据、WRF模式运行以及WRF模式spin-up留出足够时间。最外层区域采用分析Nudging四维同化,将时间分辨率为3h的GFS全球预报场同化进入WRF模式中,使得外强迫场在模式运行过程中保持准确。将最内层网格的预报结果双线性插值到观测塔所在位置,作为模式直接预报的输出量。采用的物理方案见表1。

选取的模式输出要素有6层风速(40m、70m、80m、100m、125m、150m)、3层温度(地表、40m、80m)、大气压强、地表感热通量、相对湿度、摩擦速度、边界层高度、Richardson数等共15种要素。模

式试验的时间和分辨率与观测数据相对应。

延长风电场风机的轮毂高度为70m,因此重点关注测风塔70m高度上的观测,将其作为待预报的风速,与模式直接预报和MOS预报结果作比较,判断预报方法的优劣。

4 MOS模型的建立

利用WRF中尺度气象模式及多元回归MOS方法,预报陕西省延长风电场70m高度的风速,着重研究因子数量、训练长度对预报质量的影响。

4.1 因子选取

首先,选取可能与低层风速相关的模式输出要素作为待选因子,建立回归方程时将从这些待选因子中挑出合适的因子。待选因子由中尺度气象模式WRF给出,包括6层风速($U1$ 、 $U2$ 、 $U3$ 、 $U4$ 、 $U5$ 、 $U6$ 分别代表150m、125m、100m、80m、70m、40m),3层温度($T1$ 、 $T2$ 、 $T3$ 分别代表地表、40m、80m),地表大气压强(P)及相对湿度(rh)。此外,大气稳定度对风机轮毂高度(70m)的风速有较大影响^[12],因此待选因子还可能包括地表感热通量(fx)、摩擦速度(u^*)、边界层高度(h)和Richardson数(Ri)。这些参数同样由WRF模式计算得出。

然后,根据实测风速与待选因子的相关系数确定模型中的因子。图1为风电场实测风速与各待选因子的相关系数。可以看到,实测风速与各层风速因子相关系数较高,与摩擦速度和感热通量的相关系数也相对较大,与各层温度、边界层高度和Richardson数的相关系数都小于0.1。因此,各层风速、摩擦速度和感热通量将最有可能作为预报因子

表1 WRF采用的物理参数化方案

Table 1 Physical skemes of WRF	
物理过程	参数化方案
微物理	Lin Skeme
近地面	Monin-Obukhov (Janjic Eta) scheme
陆面	unified Noah land-surface model
边界层	MYJ scheme
长波	RRTM scheme
短波	Goddard short scheme
城市冠层	NONE
积云	Kain-Fritsch scheme(第三层不采用积云方案)

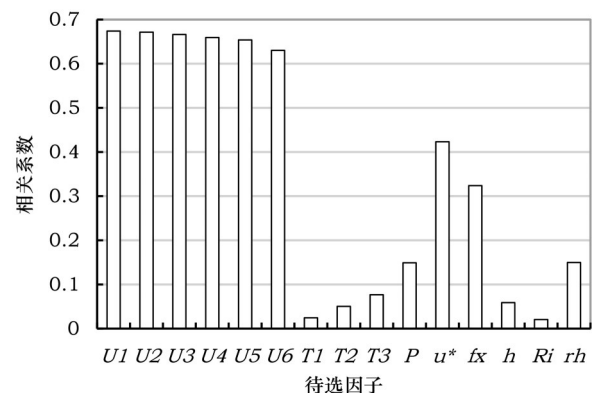


图1 实测风速与各待选因子的相关系数

Figure 1 Coefficient between observed wind speed and factors

2017年1月

进入预报方程。

4.2 因子数量的确定

要建立多元线性回归方程需要确定方程中的因子个数。理论上,因子数量越多,因子总的方差贡献就越大,残差平方和就越小。另一方面,随着因子数目的增加,每新增一个因子残差平方和下降的幅度也减小,而且增加因子会带来因子本身的各种随机因素,增加方程的不稳定性,反而造成预报质量的下降。因此,选择合适数量的因子加入到回归方程有助于预报误差的减小。

以11月20日00:00至27日00:00时7天为实验预报时间。根据国家能源局2011年下发的《风电场功率预测预报管理暂行办法》,一般风电预报的预报时长只需要一天,但是,考虑到预报应加上模式运算时间,观测数据回传时间和预报结果上传时间等,以预报时长为两天的预报能力作为标准进行研究。每次预报截取预报时间前一定长度的资料,作为训练时间,对MOS模型进行训练,即从模式输出资料中选择适当的因子,与同时段的观测资料建立回归方程。将预报时间内模式的输出结果 \hat{v}_{NWP} 代入该方程,得到MOS预报结果 \hat{v}_{MOS} 。预报效果用预报风速和实测风速的均方根误差RMSE表示。RMSE计算方法如下:

$$RMSE = \sqrt{\frac{\sum (\hat{v} - v)^2}{L}} \quad (7)$$

式中 L 为预报长度; \hat{v} 为预报风速; v 为实测风速,对预报时间求和。

通过以下方法进行关于因子数量的研究。在实验预报时间内进行滚动预报,训练长度和预报长度均不变。每过10分钟,重新训练MOS模型,得到新的回归方程及预报结果。每次预报算出模式直接预报的均方根误差 $RMSE_{NWP}$ 和MOS预报的均方根误差 $RMSE_{MOS}$ 。将7天内所有预报的误差平均可得到平均均方根误差。该平均均方根误差由多次预报均方根误差平均得到,所选因子由系统根据训练时间内的模式输出数据以及观测数据自适应决定,每次预报不尽相同。因此平均均方根误差只与因子数量和训练长度有关,与具体选取的因子无关。

给定训练长度,改变因子数量,重复上述实验,

便可得到因子数量与平均均方根误差之间的关系(图2)。预报因子等于0指的是没有经过多元回归处理,即单纯模式误差 $RMSE_{NWP}$ 。可以看到,加入前两个因子时,均方根误差有明显减小。若增加第三因子,预报效果改善不明显,并且随着因子数量的增加,出现了因子过多而产生的不稳定问题。可能的原因是,与风速相关较大的因子互相的相关系数也较大,使得只需要一到两个因子方差贡献就能达到一个相对较大的值,再增加因子方差贡献的变化也不大^[17]。因此,考虑到计算速度和方程稳定性,只选择两个因子建立方程进行预报。

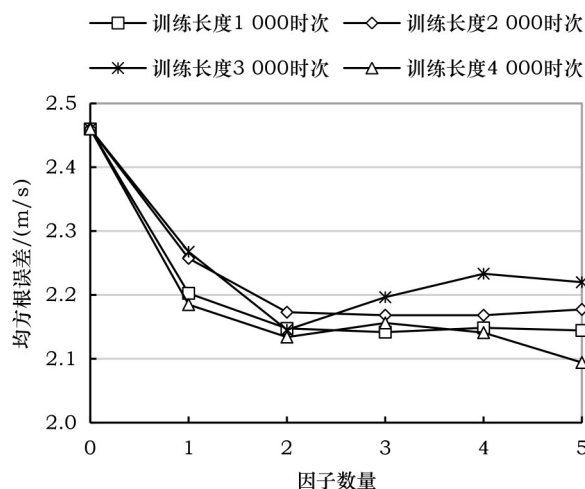


图2 平均均方根误差与因子数量的关系

Figure 2 Relationship between RMSE and number of factors

4.3 训练长度的分析

统计预报方法都需要一定长度以上的历史资料进行训练才能得到较好的预报结果。本节分析了训练长度与预报效果的关系,找出可以使用该MOS方法的最小训练长度。根据这个结果,可以依据现有的观测资料初步判断某个风电场能否应用该方法进行预报。

以11月20日00:00至27日00:00作为实验预报时间,每次预报两天。每次预报截取不同的训练长度,从300时次开始,每增加50时次计算一次,固定因子数量为2。对实验预报时间进行滚动预报,每10分钟预报一次。分别计算模式直接预报的均方根误差 $RMSE_{NWP}$ 以及经过MOS预报的均方根误差 $RMSE_{MOS}$,两者相减并在所有预报中做平均得到误差改善状况与训练长度的关系(图3)。这样做可

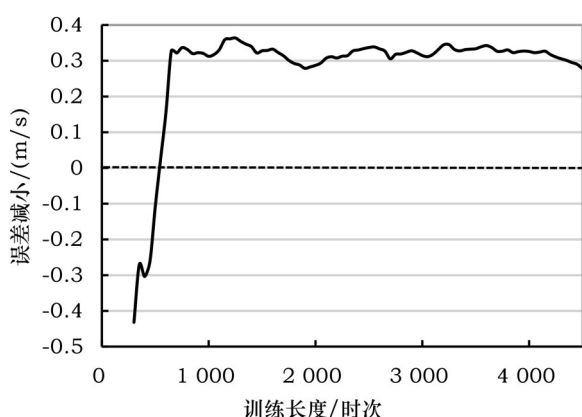


图3 误差改善与训练长度的关系

Figure 3 Relationship between RMSE improvement and training length

以去掉中尺度模式本身的误差,只研究该MOS系统的效果好坏。数值大于0,表示经MOS处理后误差变小;数值越大,MOS处理的效果越好。图3中可见,训练长度约600个点(约4天)时,预测结果有明显转折。训练长度小于该转折时,MOS效果不理想,MOS误差比模式直接输出的误差要大。训练长度大于该转折时,MOS误差比模式误差小,差距普遍在0.3m/s左右。为保证预报质量,推荐在历史资料多于一周的情况下使用该方法进行订正。

5 预报效果

5.1 30天预报效果

为了检验MOS方法的预报效果,从10月28日00:00至11月27日00:00进行为期30天的预报。规定训练长度为7天,因子数量为2个,每24h滚动预报,一次进行48h的预报。MOS的预报方程为:

$$\tilde{v}_{MOS} = b_0 + b_1 x_1 + b_2 x_2 \quad (8)$$

式中 \tilde{v}_{MOS} 为MOS预报风速; x_1, x_2 为所选取的因子; b_0, b_1, b_2 为拟合系数。选取的因子与拟合系数利用预报前7天的观测数据与模式输出数据,每次预报得到的因子及拟合系数都不尽相同。MOS可以根据模式近期的预报表现及其与观测资料的差异调整预报方程。

图4是10月28日00:00至11月27日00:00共30天70m风速的观测、模式直接预报与MOS预报数据。模式直接预报与MOS预报的误差对比如表2。

无论从均方根误差、相对误差、平均绝对误差或者相关系数来看,MOS预报都比模式直接预报有

表2 模式误差与MOS误差对比

Table 2 Comparison of WRF output and MOS

	均方根误差 /(m/s)	相对误差 /%	平均绝对 偏差/(m/s)	相关系数
NWP	2.351	42.1	1.845	0.652
MOS	2.118	39.6	1.691	0.647

所改进。其中均方根误差减少了0.233m/s。

5.2 预报实例分析

下面以几个预报实例说明该MOS方法在实际预报中的操作过程及在不同风况下的应用效果。

(1)图5a为2011年11月19日00:00至11月21日00:00的风速观测值,模式直接预报值以及MOS预报的风速预报值。MOS方法的训练时间为11月12日00:00至19日00:00,共7天。训练步骤如下:

利用11月12日00:00至19日00:00,15个待选因子的模式输出数据,根据2.2中的方法分别计算它们对实测风速的方差贡献。算得模式输出80m风速方差贡献最大并通过显著性检验,因此引入其为第一因子。在有80m风速作为第一因子的前提下,计算剩余待选因子的方差贡献。算得气压 P 方差贡献最大并通过显著性检验,引入其为第二因子。建立两个因子与实测风速的回归方程:

$$\tilde{v} = 0.678 \times v_{80} - 0.002 \times P - 184.16 \quad (9)$$

式中 \tilde{v} 为预测70m高度风速; v_{80} 为模式输出80m高度的风速; P 为气压。然后将11月19日00:00至21日00:00的模式预报80m风速以及气压代入方程中,得到MOS方法预报的70m风速。如图5a,可以看到观测风速从19日00:00至20日12:00都较小,最大不超过6m/s,从20日12:00开始增大,至大约18:00到达最大值约8m/s。而WRF模式的直接预报,在19日10:00至16:00有一个虚假的峰值,在19日18:00至20日03:00又有低估,而从20日11:00起提前估计了风速增大的时间,并高估了风速增大的强度。MOS结果对模式19日11:00起的高估有微小的调整,对中间模式低估部分有很好的补充,而最后模式高估的部分MOS也相应降低了预报值,使其与观测资料更接近。但可以看出,MOS虽然能将模式预报的系统误差有效地降低,但当模式出现明显的相位误差,如19日12:00左右的虚假大值时,MOS不能有效地消除相位误差。

2017年1月

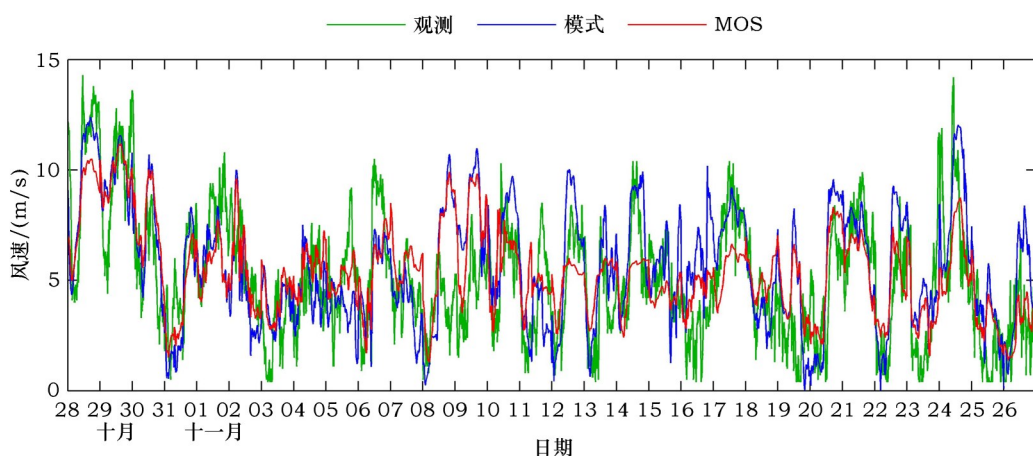


图4 10月28日00:00至11月27日00:00的观测风速,模式预报风速和MOS预报风速对比

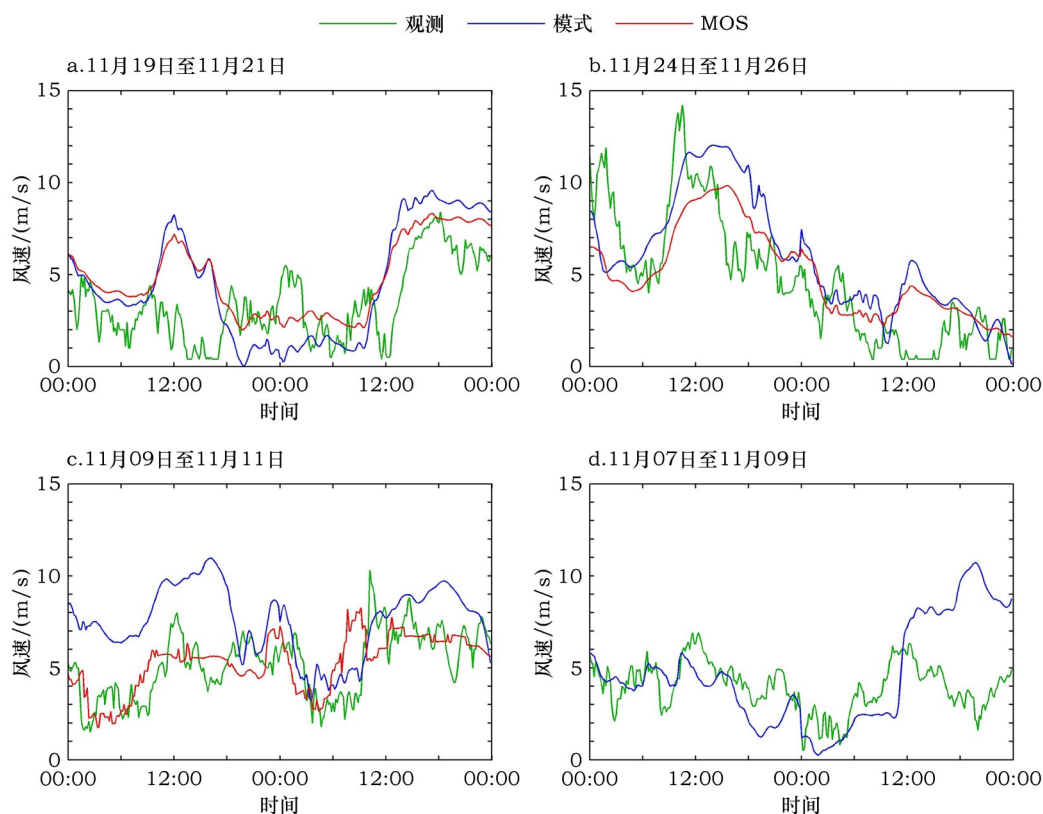
Figure 4 Comparison of observed wind speed, model output and MOS output from 00:00 Oct. 28th to 00:00 Nov. 27th

图5 观测风速,模式预报风速和MOS预报风速对比

Figure 5 Comparison of observed wind speed, model output and MOS output

从19日00:00至21日00:00,模式预报的均方根误差 $RMSE_{NWP} = 2.832$ m/s,而MOS的均方根误差 $RMSE_{MOS} = 2.398$ m/s,误差减小了0.436m/s,相对减小了15.34%。

(2)该MOS系统对大风条件下的预报结果也有所改进。图5b给出另一个例子,时间为2011年11

月24日00:00至26日00:00。可以看到,24日00:00至03:00实测风速较大,为(8~12)m/s,回落到5m/s后再次增大,于11:00左右到达最大值约为14m/s,随后逐渐回落,至25日12:00降为0。模式预报出了大风过程,但除了24日01:00至04:00和10:00至11:00的大风之外,模式整体高估了风速,

尤其24日16:00观测风速开始减小,而模式预报的风速仍约为11m/s的大值。MOS预报整体减小了风速,进而减小过大的大风估计造成的误差。

MOS的训练时间为2011年11月17日00:00至24日00:00,共7天,得到的预报方程为:

$$\tilde{v} = 0.546 \times v_{150} - 0.157 \times T_{80} - 1.850 \quad (10)$$

式中 v_{150} 为150m风速; T_{80} 为80m温度。模式直接预报的误差为2.693m/s, MOS方法的误差减小为2.312m/s, 误差减小了0.380m/s, 误差相对减小了14.14%。

(3)当模式预报发生较大偏差的时候,该MOS方法也有一定的纠正模式偏差的能力。图5c给出了2011年11月9日00:00至11日00:00的观测风速,模式预报风速以及MOS预报风速。可以看到,模式直接预报与实测资料出现了较大的偏差。11月9日00:00至19:00,模式直接预报的风速比实测风速严重偏大。图5d给出了11月7日00:00至9日00:00的观测风速与模式直接预报风速,可以看出,模式的严重偏大预报是从8日12:00左右开始的,并与实测风速呈现出相反的变化趋势。MOS系统检测到了模式直接预报风速的偏差,因此在接下来11月9日00:00至11月11日00:00的预报中,并没有选择风速作为输入因子,而选择了感热通量 f_x 和边界层高度 h 作为预报因子,得到的预报方程为:

$$\hat{v} = -0.021 \times f_x + 0.003 \times h + 4.209 \quad (11)$$

式中 f_x 为感热通量; h 为边界层高度。该预报时段内模式预报的误差为3.112m/s, MOS方法的误差为1.625m/s, 相对误差减小了47.78%。

以上3个个例中,MOS预报方程选取的因子之所以不同,是因为3个个例时间不同,对应的天气背景和模式表现也不相同。图5a是一个天气和模式表现都比较正常的状况,预报方程选择的因子就是风速和气压,这应该代表了大多数情况;结合该风电场所处的地点(中国西北部)以及研究时间(秋冬季),大风往往伴随着的是冷空气南下,气温的下降,因此风速较大的图5b, MOS选取的因子变为风速和温度;而当模式表现不佳时,MOS预报方程选取的因子变成了感热通量和边界层高度(见图5c、图5d),这两个参量恰恰反映了大气边界层垂直结构和近地面湍流的重要特征。总体来说,MOS预报

方程中所选取的因子,与近期WRF模式模拟效果关系十分密切。方程中自动选择的因子,除风速本身外,从常规的地面气压到高层(80m)温度,再到感热通量和边界层高度,这些因子从实际观测的角度来看,越来越难以获得,特别是当WRF风速模拟偏差较大时,需要采用反映大气边界层结构和湍流的因子。如果固定某些因子建立预报方程,MOS预报可能在常规状况下有较好的表现,但在某些特殊情况,如大风过程或模式预报产生较大偏差时无法得到最佳的预报。

综上所述,MOS预报建立在WRF模式预报的基础之上,与模式直接预报相比结果有所改进。一般情况下,MOS较难改变模式预报的大趋势,但根据模式近期的表现可以对模式的输出进行调整,改善预报结果。

6 结论与讨论

利用WRF模式输出的陕西省延长县风电场的多层风速、气温、边界层参量等15个要素作为因子,与实测风速构建多元回归方程,开展预报实验,通过研究和对比分析不同情况下的预报结果,得到以下规律:

(1)该方法对风速的预报比WRF模式直接输出的预报有明显提高,均方根误差减小约0.23 m/s。

(2)构建回归方程时,前两次加入因子可使误差明显减小,当加入更多因子时预报误差改善并不明显,因此以挑选两个因子进入预报方程为宜。

(3)运用多元线性回归方法进行模式后处理需要一定长度的历史数据进行训练,该最小训练长度约为600个时次。只有当现有数据长度大于最小训练长度时才可以使用该方法进行预报。

由于只采用了一个风电场的数据,因此所得结论可能有一定局限性,对于其他的风电场,可能需要不同的因子数量和训练长度。另外,本文只考虑了风速的预报,而没有考虑风向的预报,不同的风向由于涉及地形、尾流等原因,对发电量也有较大影响。另外,由于实际风速与各预报因子的关系可能与当前季节与天气形势有关,若有足够长的历史资料和模式反算数据,对季节和天气型进行分类,建立不同条件下的模型,应能得到更好的预报结果。

2017年1月

致谢:本文第一作者感谢中国气象局风能太阳能资源中心宋丽莉研究员的指导和帮助,感谢中国科学院大气物理研究所徐晶晶博士为本文提供了观测数据。

参考文献(References):

- [1] 范高锋,裴哲义,辛耀中. 风电功率预测的发展现状与展望[J]. 中国电力, 2011, 44(6): 38-41. [Fan G F, Pei Z Y, Xin Y Z. Wind power prediction achievement and prospect[J]. *Electric Power*, 2011, 44(6): 38-41.]
- [2] Boone A. Simulation of Short-Term Wind Speed Forecast Errors Using a Multi-Variate Arma (1, 1) Time-Series Model[D]. Stockholm: Royal Institute of Technology, 2005.
- [3] Milligan M, Schwartz M N, Wan Y. Statistical Wind Power Forecasting for U.S. Wind Farms[C]. Seattle: 17th Conference on Probability and Statistics in the Atmospheric Sciences/2004 American Meteorological Society Annual Meeting, 2003.
- [4] Louka P, Galanis G, Siebert N, et al. Improvements in wind speed forecasts for wind power prediction purposes using kalman filtering[J]. *Journal of Wind Engineering and Industrial Aerodynamics*, 2008, 96(12): 2348-2362.
- [5] Mohandes M A, Halawani T O, Rehman S, et al. Support vector machines for wind speed prediction[J]. *Renewable Energy*, 2004, 29(6): 939-947.
- [6] 戚双斌,王维庆,张新燕. 基于SVM的风速风功率预测模型[J]. 可再生能源, 2010, 28(4): 25-28. [Qi S B, Wang W Q, Zhang X Y. Modeling buiding for wind speed and wind power prediction based on SVM[J]. *Renewable Energy Resoures*, 2010, 28(4): 25-28.]
- [7] Alexiadis M C, Dokopoulos P S, Sahsamanoglou H S, et al. Short-term forecasting of wind speed and related electrical power[J]. *Solar Energy*, 1998, 63(1): 61-68.
- [8] 李洪涛. 人工神经网络和模糊技术在风能预测中的应用[J]. 中国电力, 2012, 45(8): 69-73. [Li H T. Application of artificial neural network and fuzzy logic on wind power prediction[J]. *Electric Power*, 2012, 45(8): 69-73.]
- [9] Landberg L, Giebel G, Nielsen H A, et al. Short-term prediction-an overview[J]. *Wind Energy*, 2003, 6(3): 273-280.
- [10] 吴琼,贺志明,聂秋生,等. 动力降尺度法对鄱阳湖区风能资源模拟效果分析[J]. 资源科学, 2012, 34(12): 2337-2346. [Wu Q, He Z M, Nie Q S, et al. Evaluation of wind energy simulated by dynamical downscaling methods for Poyang Lake[J]. *Resources Science*, 2012, 34(12): 2337-2346.]
- [11] 王勇,李照荣,李晓霞,等. 风电功率预报方法研究进展[J]. 干旱气象, 2011, 29(2): 156-160. [Wang Y, Li Z R, Li X X, et al. Forecast method advanced on wind electricity[J]. *Journal of Arid Meteorology*, 2011, 29(2): 156-160.]
- [12] Giebel G, Kariniotakis G, Halliday J, et al. Next Generation Short-Term Forecasting of Wind Power- Overview of the ANEMOS Project[C]. Athens: European Wind Energy Conference, 2006.
- [13] 徐晶晶,胡非,肖子牛,等. 风能模式预报的相似误差订正[J]. 应用气象学报, 2013, 24(6): 731-740. [Xu J J, Hu F, Xiao Z N, et al. Analog bias correlation of numerical model on wind power prediction[J]. *Journal of Applied Meteorology Science*, 2013, 24(6): 731-740.]
- [14] 傅荣,王维庆,何桂雄. 基于气象因子的BP神经网络风电场风速预测[J]. 可再生能源, 2009, 27(5): 86-89. [Fu R, Wang W Q, He G X. The forecasting of wind speed in wind farm based on the meteorological factors with BP neural network[J]. *Renewable and Sustainable Energy Reviews*, 2009, 27(5): 86-89.]
- [15] 程兴宏,陶树旺,魏磊,等. 基于WRF模式和自适应偏最小二乘回归法的风能预报试验研究[J]. 高原气象, 2012, 31(5): 1461-1469. [Cheng X H, Tao S W, Duan W. Short-term wind power forecasting experiment based on WRF model and adapting partial least square regression method[J]. *Plateau Meteorology*, 2012, 31(5): 1461-1469.]
- [16] Glahn H R, Lowry D A. The use of model output statistics in objective weather forecasting[J]. *Journal of Applied Meteorology*, 1972, 11(8): 1203-1211.
- [17] Landberg L. A mathematical look at a physical power prediction model[J]. *Wind Energy*, 1998, 1(1): 23-28.
- [18] 刘还珠,赵声蓉,陆志善,等. 国家气象中心气象要素的客观预报-MOS系统[J]. 应用气象学报, 2004, 15(2): 181-192. [Liu H Z, Zhao S R, Lu Z S, et al. Objective element forecastsat Nmc-a MOS system[J]. *Journal of Applied Meteorology Science*, 2004, 15(2): 181-192.]
- [19] 陈豫英,陈晓光,马金仁,等. 风的精细化MOS预报方法研究[J]. 气象科学, 2006, 26(2): 210-215. [Chen Y Y, Chen X G, Ma J R, et al. A study on subtle MOS forecasting method of wind[J]. *Scientia Meteorologica Sinica*, 2006, 26(2): 210-215.]
- [20] 林行,马艳,杨育强,等. 动态统计预报模型及其在2008奥帆赛风场预报中的应用[J]. 气象科学, 2008, 34(S1): 199-205. [Lin X, Ma Y, Yang Y Q, et al. Dynamical statistical forecasting model and its application on wind forecast of 2008 Olympic Sailing Race [J]. *Meteorological Monthly*, 2008, 34(S1): 199-205.]
- [21] Ranaboldo M. Multiple Linear Regression Mos for Short-Term Wind Power Forecast[D]. Barcelona: University of Barcelona, 2011.
- [22] 黄嘉佑. 气象统计分析与预报方法[M]. 北京: 气象出版社, 2004. [Huang J Y. Methods of Meteorological Statistics Analysis and Predition[M]. Beijing: China Meteorological Press, 2004.]

Model output statistics and wind power numerical prediction

XIAO Qingyao^{1,2}, HU Fei¹, FAN Shaojia³

(1. LAPC, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China;

3. School of Atmospheric Sciences, Sun Yat-sen University, Guangzhou 510275, China)

Abstract: Instead of using the direct output from numerical weather prediction models, model post-processing using available observations are needed for meeting the rigorous requirements of wind speed prediction at wind farms. Here, a model output statistics (MOS) model is established using suitable factors within wind speed and temperature at different heights in the boundary layer, friction velocity, intensity of turbulence, Richardson number, air pressure and other meteorological elements from the mesoscale numerical weather prediction system WRF model output. Model fitting, selected factors and historical observation data in the same period with multi-linear regression formulas, and predicting wind speed at a wind farm in Shanxi, China was done. Effects on the prediction quality of a number of factors selected by the model and training length of the model are researched in detail. We found that generally there is no need for a large number of factors and the recommended number is two. Historical data with length of at least 600 are required for the quality of MOS prediction to be improved in evidence. Establishing the regressive formula, meeting the two conditions above, the root mean square error of MOS prediction would be about 0.23 m/s smaller compare to the WRF direct output. Considering the short calculation time and easy realization of this method it can be broadly applied to wind power prediction.

Key words: MOS; multi-regression; wind power prediction; WRF